

Nr.	Hypotese	(A) Grafisk	(B) Numerisk	(C) Uddybende
-	Modellen	$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + e_i$ <p>Opstil evt. m. opgaverelevante termer, eller omdøb variable, eks.: løn = x_1, Uddannelse = x_2 og profit = y</p>		
-	Std.forudsætninger	<p>I) $E(e_i) = 0$ II) $Var(e_i) = \sigma^2$ for $i = 1, \dots, n$. III) $Cov(e_i, e_j) = 0$ for $i \neq j \Rightarrow$ Uafhængige restled</p> $\Rightarrow e_i \sim Nf(0, \sigma^2) \dots \text{IV) hvid støj}$		
-	Standardiserede residualer	$e_i^* = \frac{e_i}{\sqrt{Var(e_i)}}$ <p>hvor $var(e_i)$ afhænger af værdien af den forklarende variabel, mens $var(e_i^*)$ er konstant (homoskedasticitet) for alle $i = 1, \dots, n$, så de kan sammenlignes direkte!</p>		
1	Ad. I) Linearitet	<ul style="list-style-type: none"> - Plot (y, \hat{y}) -dia.: skal ligge om 45° linie. - Plot (x_i, e_i^*) – kontroller systematik 	$R^2 = \frac{SAK_y - SRK}{SAK_y} > 0,5 \Rightarrow$ relativt høj	$R^2 =$ Determinationskoefficient, angiver den andel af den totale variation i responsvariablen (y), der kan beskrives ud fra den lineære relation til de forklarende variabel (x_i)
2	Ad. IV) Normalfordelte std. residualer	Plot fraktil-dia. (Prob Plot): • Skal ligge tilfældigt om en ret linie		
3	Ad. II) Homoskedasticitet	- Plot (x_i, e_i^*) : skal ligge usystematisk i et bredt bånd, uden for mange punkter uden for $(x_j, 2,0)$		
4	Ad. III) Uafhængige restled (Autokorr.-kontrol) NB: Kun ved tidsrække obs.	- Plot (t, e_i^*) : Kommenter passede antal fortegnsskift*	$DW = \frac{\sum_{i=2}^n (\hat{e}_i - \hat{e}_{i-1})^2}{\sum_{i=1}^n \hat{e}_i^2}$ <p>0 d₁ d₂ d₃ d₄ 4 PA NA ID IA ID NA PA/NA=Pos./Neg. autokorr. ID=Indecisionsområdet IA = Ingen autokorrelati</p> <p>Forklar hvorfor PA/NA (hvis intuitivt)</p>	- $(t, e_i^*) =$ bløde kurver \Rightarrow pos. autokorr. - $(t, e_i^*) =$ ZIG-ZAG \Rightarrow neg. autokorr. - DW kontrollerer for AR(1)-korrelation: dvs., om $e_i = \rho e_{i-1} + v_i,$ hvor v_i er uaf. NF restled \square $NF(0, \sigma^2)$
5	Betydning af enkelt-obs.	- Plot (x_j, y) -dia.: • Kommenter om nogle obs. ligger "isoleret" og dermed påvirker estimationen meget • Kommenter hvorledes de påvirker parameter estimaterne.	Cook's afstand: $D_i = \frac{(\hat{\beta} - \hat{\beta}_{(i)})^T (\bar{X}^T \bar{X}) (\hat{\beta} - \hat{\beta}_{(i)})}{(p+1)s^2}$, $\hat{\beta}_{(i)}$ er MK parameterestimatet uden obs. nr. i.	- $D_i =$ forskellen i parameterestimat med og uden obs. nr. i. - Anvendes til at udpege obs., der bør underkastes nærmere undersøgelse 1. Udpeg størst Cook's D. 2. Sig hvorfor: Stor residual, ekstrem x-kombination eller begge. 3. Kommenter $M_i (=$ Mahalanobis): intra- vs. ekstrapolation/yderkant af obs.
6	Multikollinearitet:	- Plot (x_{1i}, x_{2i}) og kontroller for systematik	$\rho = \frac{cov(X_1, X_2)}{\sqrt{var(X_1) \cdot var(X_2)}} \text{ (viser kun lineær korr.)}$	Def: der eksisterer et eller flere lineære bånd af typen: $c_0 + c_1 x_{1i} + \dots + c_p x_{pi} = 0$ Konsekvens: Stor $Var(\beta_i) \Rightarrow$ fortolkningsproblemer af parametre /effekten af enkelte forklarende variable på responsvariablen og brede konf/pred. Symptomer: 1) Lave værdier af partielle t-str. ses med stor R^2 . 2) Parameterskøn har ikke forventet fortegn eller er insignifikante imod øko.teori. 3) Udeladelse af enkelt obs. eller enkelte forklarende variable ændrer parameterskøn meget.
7	$H_0: \beta_1 = \beta_2 = \dots = \beta_p$		$V = \frac{R^2 / p}{(1 - R^2) / (n - p - 1)} \square F_{0,95}(n, n - p - 1)$	
8	$H_0: \beta_i = \beta_{i0}$ (kan også være tosiddet)		$T_i = \frac{\hat{\beta}_i - \beta_{i0}}{\sqrt{var(\hat{\beta}_i)}} \square t(n - p - 1)$	